A Boolean Model of Molecular Signaling Pathways

Olivier Niel

A Boolean Model of Molecular Signaling Pathways

Olivier RP Niel [1]

Department of Biology - Université de Nice Sophia Antipolis, Parc Valrose, 28 avenue de Valrose, F06000 NICE, France

Faculté de Médecine de Nice - Université de Nice Sophia Antipolis, 28 avenue de Valombrose, F06107 NICE, France

(1) E-mail: o.r.p.niel@free.fr

Abstract

The usual representation of signaling pathways in molecular biology consists of diagrams. These diagrams are often unclear and confusing, thus making global understanding, prediction and computerization of the signaling pathways impossible. Here we propose a novel representation of signaling pathways, based on a mathematical model. This model associates a simple equation with any signaling pathway. We used this model to study several pathways, such as Akt, leading to strong predictions and powerful *in silico* experiments.

Keywords: Boolean, signaling pathway, Akt, set theory, bioinformatics

From the CPBT-08 meeting, Toulon, France.

1. Introduction

The amount of data concerning signaling pathways in molecular biology has increased dramatically in the past few years. These data are usually analyzed and presented as diagrams, in which each protein of a signaling pathway is linked to its target protein by an arrow. Unfortunately, this method is very limited. Indeed, these diagrams become very confusing with even a small increase in the number of proteins involved. Moreover, because of their graphic nature, they don't lead to any calculation, prediction or computerization. Here we propose a novel mathematical model of signaling pathways, based on the Zermelo-Fraenkel with axiom of Choice Set Theory [1,2] and on Boolean functions [3,4]. This model associates a simple equation with any signaling pathway. It is then easy to analyze the behavior of a given pathway under various conditions, to predict the way a new protein will interact with an existing pathway, to verify the logic of an hypothesis even before any actual experiment, and to predict, by computerization, the behavior of a several signaling pathways at the same time.

2. Methods

Let E^k be a *n*-element binary set according to the Zermelo-Fraenkel Set Theory with the axiom of choice (ZFC) [1,2]. E^k contains elementary molecular parameters (meaning fundamental parameters) of any living organism of choice. Let *k* be defined in $\mathbb{R}$. Then:

$$E^k = \{a_1, \ldots, a_n\} / \forall m \in \mathbb{R}, (a_m = 0 \vee a_m = 1)$$

Let's define *f* as a linear Boolean function [3,4] as follow: $f : E^k \to \{0,1\}$. That application can be associated with its own truth table.

We will use Karnaugh maps and Gray code to associate equations to signaling pathways.

3. Results

3.1 The theoretical approach

Let's modelize a pathway. The key protein of this pathway is ${}^k a_m^i$. Four different proteins are directly interacting with ${}^k a_m^i$: ${}^k a_x^i$, ${}^k a_y^i$, ${}^k a_z^i$ and ${}^k a_w^i$. We'd like to obtain a simple equation that describes the activity of ${}^k a_m^i$, instead of just having a graphic representation of these five proteins. The usual diagram representing the pathway would be the following.

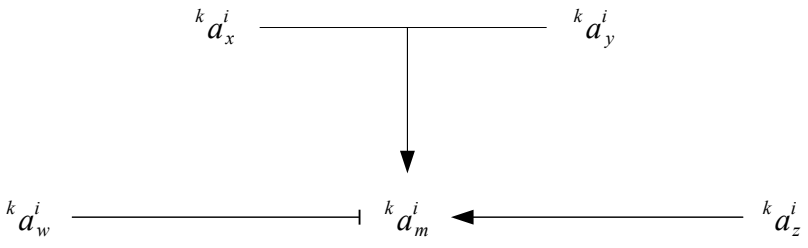

Figure 1: Diagram of the ${}^k a_m^i$ pathway

The corresponding Karnaugh map is the following. It represents all the possible activation states of the proteins that directly interact with ${}^{k}a_{m}^{i}$. The Karnaugh map is directly derived from actual gain-and-loss-of-function experiments that led to the diagram above.

${}^{k}a_{m}^{i}$ value	$\overline{{}^{k}a_{w}^{i}}$ $\overline{{}^{k}a_{x}^{i}}$	$\overline{{}^{k}a_{w}^{i}}$ ${}^{k}a_{x}^{i}$	${}^{k}a_{w}^{i}$ ${}^{k}a_{x}^{i}$	${}^{k}a_{w}^{i}$ $\overline{{}^{k}a_{x}^{i}}$
$\overline{{}^{k}a_{y}^{i}}$ $\overline{{}^{k}a_{z}^{i}}$	0	0	0	0
$\overline{{}^{k}a_{y}^{i}}$ ${}^{k}a_{z}^{i}$	1	1	0	0
${}^{k}a_{y}^{i}$ ${}^{k}a_{z}^{i}$	1	1	0	0
${}^{k}a_{y}^{i}$ $\overline{{}^{k}a_{z}^{i}}$	0	1	0	0

Figure 2: Karnaugh map of the ${}^{k}a_{m}^{i}$ *pathway*

From this Karnaugh map, it is straightforward to deduce the Boolean equation that describes the activity of ${}^{k}a_{m}^{i}$ as follow.

After factorization, we have the following Boolean equation: ${}^{k}a_{m}^{i}=\overline{({}^{k}a_{w}^{i})}({}^{k}a_{z}^{i}+{}^{k}a_{x}^{i}.{}^{k}a_{y}^{i})$

This equation describes the activity of ${}^{k}a_{m}^{i}$ according to the activity of the four known proteins that directly interact with ${}^{k}a_{m}^{i}$.

The advantages of using this equation instead of the diagram shown above to describe the ${}^{k}a_{m}^{i}$ pathway will be discussed below.

3.2 An example: the Akt signaling pathway

The Akt signaling pathway [5,6] is involved in many cellular responses. Its role remain unclear, but seems to be of a major importance in many pathologies, like type II diabetes and cancer [7,8]. Let's modelize the Akt effects on the protein Bad.

In the beginning of 2005, the knowledge of the Akt effects on Bad during apoptosis was represented as follow [9].

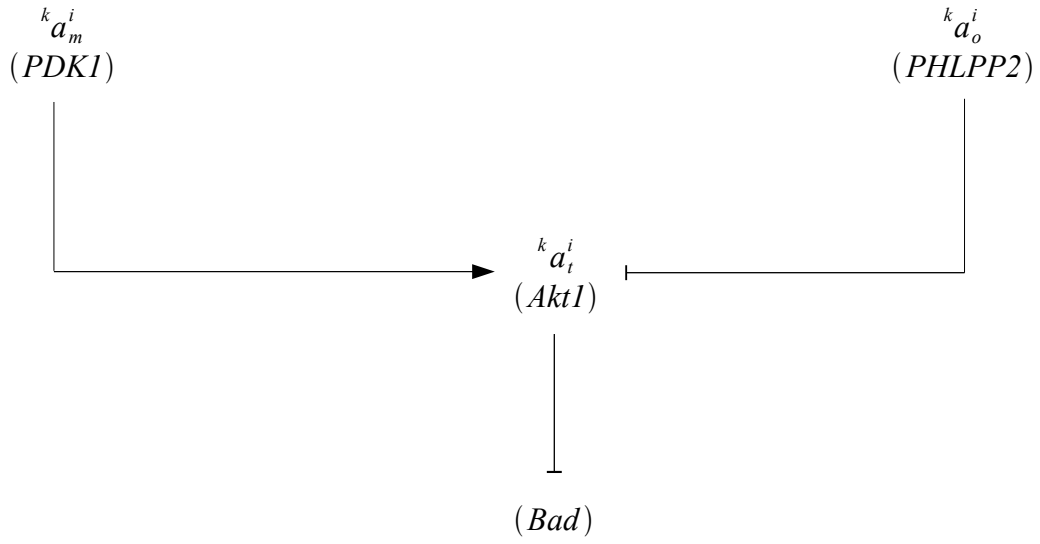

Figure 3: Diagram of the Akt pathway - preliminary

The corresponding Karnaugh map, based on the previous diagram and on data from the literature back in 2005 is shown below.

${}^{k}a_{t}^{i}$ value	$\overline{{}^{k}a_{m}^{i}}$	${}^{k}a_{m}^{i}$
$\overline{{}^{k}a_{o}^{i}}$	0	1 or 0
${}^{k}a_{o}^{i}$	0	0

Figure 4: Karnaugh map of the Akt pathway - preliminary

The fact that the ${}^{k}a_{t}^{i}$ value, when ${}^{k}a_{m}^{i}$ and $\overline{{}^{k}a_{o}^{i}}$ is both 0 and 1at the same time means that a factor is still unknown, but plays a role in this pathway. Let's try to modelize this fact.

The unknown factor should account for an activation of ${}^{k}a_{t}^{i}$ in synergy with ${}^{k}a_{m}^{i}$.
Thus, we can modelize the missing factor as shown in the diagram below.

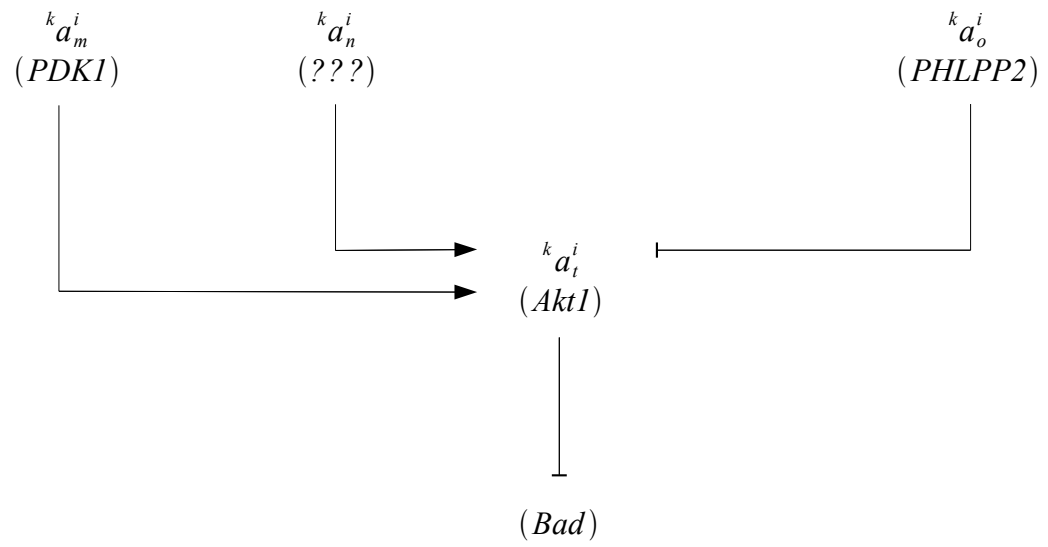

Figure 5: Diagram of the Akt pathway – final

The corresponding Karnaugh map is the following.

${}^{k}a_{t}^{i}$ value	$\overline{{}^{k}a_{m}^{i}}$ $\overline{{}^{k}a_{n}^{i}}$	$\overline{{}^{k}a_{m}^{i}}$ ${}^{k}a_{n}^{i}$	${}^{k}a_{m}^{i}$ ${}^{k}a_{n}^{i}$	${}^{k}a_{m}^{i}$ $\overline{{}^{k}a_{n}^{i}}$
$\overline{{}^{k}a_{o}^{i}}$	0	0	1	0
${}^{k}a_{o}^{i}$	0	0	0	0

Figure 6: Karnaugh map of the Akt pathway - final

From this Karnaugh map, it is straightforward to deduce the Boolean equation that describes the activity of Akt and its effects on Bad: ${}^{k}a_{t}^{i}=({}^{k}a_{m}^{i}.{}^{k}a_{n}^{i}.\overline{{}^{k}a_{o}^{i}})$.

Some experiments conducted later [10] in 2005 showed that this reasoning was correct. ${}^{k}a_{n}^{i}$ was identified as mTORC2.

4. Discussion

4.1 Advantages of the Boolean model

The first advantage of using a Boolean equation instead of a diagram to describe a pathway is a better understanding of the pathway itself. Diagrams quickly become confusing, even with a small increase in the number of proteins involved, whereas an equation is always easy to understand and to test.

The second advantage of our model is the possibility to predict the activity of a protein without doing any actual experiment. Indeed, let's consider two different pathways in the same cell. Let's admit that pathway 1 is described by the equation ${}^k a^i_{m1} = (\overline{{}^k a^i_z})({}^k a^i_x + {}^k a^i_y)$ whereas pathway 2 is described by the equation ${}^k a^i_{m2} = ({}^k a^i_w)({}^k a^i_z)$. Let's assume that previous experiments have shown that ${}^k a^i_{m1} = 1$ and ${}^k a^i_w = 1$. It is then straightforward to predict that ${}^k a^i_{m2} = 0$, as $(\overline{{}^k a^i_z}) = 1 \Rightarrow {}^k a^i_z = 0 \Rightarrow {}^k a^i_{m2} = 0$. That kind of prediction is much more complicated, if not impossible to do with a diagram.

The third advantage of our Boolean model is the possibility to use it to computerize one or several signaling pathways at the same time, leading to the modelization of entire cellular functions. Indeed, because of the dramatic increase in the number of signaling pathways described each year, the use of bioinformatics is now mandatory in order to provide a global analysis of a cellular process. Since any programming language (C/C++, PHP, Java, Python, Caml,...) provides native instructions for Boolean programming, it is particularly easy to use our model to perform powerful *in silico* experiments. Such a global approach is impossible to achieve with a graphic representation of signaling pathways, like diagrams.

The fourth advantage of our model is its possibility to facilitate a multidisciplinary approach in molecular biology. Indeed, by expressing the results of biological processes as simple equations, our model can be used by biologists, but also by mathematicians or physicists. The study of a signaling pathway can therefore benefit from the power of a mathematical approach (equations solving, application of theorems).

4.2 Limits of the Boolean model

The first limit might be the mathematical knowledge required to understand and use the model, even if Boolean equations and functions are really simple to handle.

The second limit is inherent to any modelization: if the knowledge of a process is too limited, then its modelization can lead to wrong results. Mathematical models are very powerful tools, but they don't spare the trouble of doing actual experiments.

5. Conclusion

We introduced a novel approach in the analysis of signaling pathways in molecular biology. We used a mathematical model, based on the ZFC Set Theory and on Boolean functions, to associate a simple equation with any given signaling pathway. By doing so, we were able to significantly improve data analysis. Our model leads to a better global understanding of signaling pathways, along with great predictions, powerful in silico experiments, and multidisciplinary approaches. We expect this Boolean model to be widely used amongst biologists, mathematicians and physicists, in an attempt to understand molecular biology better.

Acknowledgments

We'd like to thank LNU, JSN and JMP for insightful talks and reviewing of this paper.
We'd also like to thank the CPBT-08 for publication of this work.
No conflict of interest.

Online pedagogic resources

Karnaugh map (http://en.wikipedia.org/wiki/Karnaugh_map)
Boolean functions (http://en.wikipedia.org/wiki/Boolean_function)
ZFC Set Theory (http://en.wikipedia.org/wiki/Zermelo%E2%80%93Fraenkel_set_theory)

References

1. Krivine, JL. in *Theorie des ensembles*. Cassini, Paris (1998).

2. Jech, T. in *Set theory*. Academic Press, New York (1978).

3. Koppelberg, S. in *General Theory of Boolean algebras*. Handbook of Boolean Algebras, Amsterdam (1989).

4. Sikorsky, R. in *Boolean Algebra*. Springer-Verlag, Berlin (1969).

5. Duronio, V. The life of a cell: apoptosis regulation by the PI3K/PKB pathway. *Biochem J*. **415(3)**, 333-344 (2008).

6. Manning, B.D., Cantley L.C. Akt/PKB signaling: navigating downstream. *Cell*. **129(7)**, 1261-1274 (2007).

7. Hoshino, Y., Sato, E. PKB/Akt is required for the completion of meiosis in mouse oocytes. *Dev. Biol.* **314(1)**, 125-136 (2008).

8. Mendoza, M., Blenis, J. PHLPPing it off: phosphatases get in the Akt. *Molecular Cell*. **25**, 798-800 (2007).

9. Gao, T., Furnari, F., Newton, AC. PHLPP: a phosphatase that directly dephosphorylates Akt, promotes apoptosis, and suppresses tumor growth. *Molecular Cell*. 18(1), 13-24 (2005).

10. Sarbassov, DD., Guertin, DA., Ali, SM., Sabatini, DM. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science*. **307(5712)**, 1098-1101 (2005).